



Evidence Based Library and Information Practice

Evidence Summary

Enhanced Catalogue Records Positively Impact Circulation but Are Not Used to Their Potential in Patron Searching

A Review of:

Tosaka, Y., & Weng, C. (2011). Reexamining content-enriched access: Its effect on usage and discovery. *College & Research Libraries*, 72(5), 412-427.

Reviewed by:

Cari Merkley
Associate Professor
Mount Royal University Library
Calgary, Alberta, Canada
Email: cmerkley@mtroyal.ca

Received: 1 Mar. 2012

Accepted: 16 June 2012

© 2012 Merkley. This is an Open Access article distributed under the terms of the Creative Commons-Attribution-Noncommercial-Share Alike License 2.5 Canada (<http://creativecommons.org/licenses/by-nc-sa/2.5/ca/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly attributed, not used for commercial purposes, and, if transformed, the resulting work is redistributed under the same or similar license to this one.

Abstract

Objective – To determine how content-enriched catalogue records impact the circulation rates of print resources in four subject areas, and to investigate how this additional metadata influences OPAC searching and item retrieval.

Design – Analysis of circulation data, bibliographic records, and OPAC search logs.

Setting – A library at a four-year undergraduate residential college in the North-eastern United States.

Subjects – Bibliographic records for 88,538 titles; data from 7,782 circulation transactions;

and 130 OPAC search strings and related circulation data.

Methods – In the first part of the study, bibliographic records for print items published since 1990 were extracted from the library's integrated library system (ILS) in the following Library of Congress (LC) classes: D, E, F, H, J, L, P, Q, R, S, and T. It is assumed that electronic books were excluded from this study because their usage is not tracked in the ILS. These LC classes were chosen to correspond to the subject areas targeted by the researchers for comparison – "history, social sciences, language and literature, and science and technology" (p. 416). The data file included the publication date of the title, as well as values for the MARC fields identified

by the researchers as containing content-enriched data. These fields were MARC 505 (an item's table of contents or list of works included), MARC 520 (summaries or annotations), and MARC 856 (URL to electronic location of related material or electronic copy) (p. 416; Library of Congress Network Development and MARC Standards Office, 2003, 2008a, 2008b). The authors analyzed records for 88,538 titles and determined the total number of records containing each of the MARC fields either singly or in combination.

Data relating to circulation transactions for items located in these LC classes from January to May 2009 was also identified. Like the bibliographic records, circulation data was pulled for print items only. The researchers identified 7,782 circulation transactions that met the study criteria for the period in question.

In the second part of the study, circulation data for September 22, 2009 was obtained and sorted into the four subject categories identified in Part I of the study. The authors indicate that this date was chosen at random, but do not specify how. Researchers compared the records of the 133 titles borrowed that day from the LC classes studied to the OPAC search logs from September 16-22, 2009 to determine which searches led to the circulation of these items. The authors felt that searches resulting in checkouts on the day in question may have begun earlier in the week. The searches that led to borrowing were recorded and categorized as keyword, title, author, or other searches. If a user entered a title or author name into the keyword field, these were classed as known item searches in the appropriate categories. The authors identified and analyzed 130 searches relating to circulated items.

Main Results – In the first part of the study, the number of catalogue records that contained MARC 505, 520, and/or 856 fields significantly increased for titles published between 1990 and 2007, with a slight decrease in 2008. MARC 505 was the most common content-

enriched field until 2000, after which the presence of MARC 856 grew significantly. The MARC 520 field was used least often, making it difficult to draw firm conclusions about its impact on circulation.

The incidence of enhanced records was very low among older books in the study. Only 14.3% of items published between 1990 and 1994, and 19.3% of items published between 1995 and 1999, had records that contained MARC 505, 520, or 856 fields. In contrast, the percentage of enhanced records was very high (80.9%) for items published between 2005 and 2008. The authors acknowledged that these stark imbalances created skewed comparison data for items published in these date ranges. As such, they suggested that the data for titles published between 2000 and 2004 offered the most balanced comparison because the numbers of enhanced and non-enhanced records were almost equal. The overall circulation of items with enhanced records published between 2000 and 2004 was 2.9% higher than for items with non-enhanced records, constituting a relative percentage difference of 30.7%. The relative percentage difference in this period was higher for books in science and technology (36.9%), followed by history (34%), language and literature (30.6%), and social sciences (25.7%). Enhanced records also had a positive impact on circulation for items published between 1990 and 2000 over their non-enhanced counterparts; however, this positive growth levelled off for items published between 2005 and 2008, with almost equal circulation rates between items with enhanced and non-enhanced records during this period. The impact of the three MARC fields was examined, and the presence of the MARC 505 field was most associated with increased circulation rates, in part because it was the most commonly used field of the three for the period in question. The number of records with MARC 520 and 856 fields was not sufficient to draw firm conclusions about their impact on circulation. While not the focus of the study, the circulation data also suggested a preference for current titles among all four subject areas, most significantly among the social sciences and science and technology.

The second part of the study found that keyword searching was the most common strategy employed by patrons, with 49.6% of the 130 searches examined falling into this category. Keyword searches most commonly led to the borrowing of items from the history LC classes, while title searches were most common in science and technology. Known item searches (title or author) accounted for 45.9% of the overall searches analyzed. However, in most cases, the search terms used that led to a title circulating were found in the title and subject fields, rather than in a content-enriched MARC field. The researchers suggested that this may be due to the appearance of search results in the OPAC (brief rather than full record) and the way relevancy sorting was calculated, as contents notes were not given a high weighting in the OPAC's formula.

Conclusion – The study found that enhanced catalogue records led to higher circulation rates in the four subject areas studied. The increased proportion of content-enriched records in the overall catalogue in recent years suggested that their value had been recognized by the library. The limited role these enhanced fields played in the September 22, 2009 searches suggested that further work on improving how this information is displayed to users in the OPAC and sorted is needed. The researchers identified areas for future research including the role of the publication date and the impact of improvements to the display of content fields in the OPAC on the circulation of items with content-enriched records.

Commentary

This work adds to an existing argument in the library literature: that enhanced catalogue records are correlated with increased circulation. Their presentation of the study data is measured and thorough, acknowledging the disproportionate impact small sample sizes have on the percentage differences in several of their data sets and their limited ability to draw reasonable conclusions from such numbers. It is useful to

have both the absolute and relative percentage differences in circulation to better gauge the difference made by content-enriched records. Even better, it is always clear which percentage difference, absolute or relative, is the one being stated by the authors. Tosaka and Weng do acknowledge, however, that enhanced catalogue records may not be the most important factor in determining circulation rates. Publication date played a significant role in the study results, with more recent titles circulating more across all disciplines. Teasing out whether these circulation gains are due to the fact that newer records are more likely to be enhanced, or newer materials are just of more interest to users, would be a difficult but worthwhile task.

The findings of the second part of the study suggest that enhanced records or date of publication are not the only factors impacting circulation rates: how the patron and library technology interacts with them is perhaps the richer field for study moving forward. The authors write, "To achieve content-enriched access, it is necessary to have a well-designed data-mining mechanism to dig out content-enriched components to system retrieval ability and postsearch evaluation" (p. 413). It was striking how small a role the content-enhanced fields appeared to play in the searches and subsequent circulations on the day studied by the researchers. They proposed explanations for the disconnect between the findings of the first and second parts of the study raise important questions for practitioners. Why spend the energy, time, and staff dollars on enhancing catalogue records, to then only display the brief view of the record or sort by publication date in your OPAC? Is this a choice on the part of the library, or a consequence of ILS display design?

Alternatively, if publication date is one of the most important criteria for patrons in selection of material in an OPAC, is default relevancy ranking a help or an obstacle for users? While the question of to enhance or not to enhance records appears to have been largely answered (as evidenced by the prevalence of enriched records today), it is clear that the work of

creating a better and more responsive access point to our collections is never done.

References

Library of Congress Network Development and MARC Standards Office. (2003). Guidelines for the use of field 856. Retrieved February 27, 2012, from <http://www.loc.gov/marc/bibliographic/bd856.html>

Library of Congress Network Development and MARC Standards Office. (2008a). 505 - Formatted Contents Note (R). Retrieved February 27, 2012, from <http://www.loc.gov/marc/bibliographic/bd5xx.html>

Library of Congress Network Development and MARC Standards Office. (2008.). 520 - Summary, Etc. (R). Retrieved February 27, 2012, from <http://www.loc.gov/marc/bibliographic/bd520.html>